



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Fledgling data science

Citation for published version:

Campagnolo, GM 2019, 'Fledgling data science: A study of the freelance football analytics community'.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Fledgling data science: a study of the freelance football analytics community

G.M. Campagnolo (UniEd/Turing)

G. Fele (Unitn)

contacts: g.campagnolo@ed.ac.uk

Public Analyst

Scamps and student-bedroom YouTube sensations, this is a community of young students or early career professionals (mainly between 21 and 26 years old) whose side project is to populate the blogosphere and the twitter-sphere with statistically informed football commentaries to get noticed and find a day job into professional football.

Public Analyst

#OptaProForum Line-Up

Speakers:

@Torvaney

Freddie Wilson

@JohnMarkSisman

@statlurker

@penaltyblog

@FinerMargins , @AlexRathke &

@FerdiaOHanrahan

Posters:

Ben Woolcock

@dperdomomeza1 & @BobbyGardiner

Gian Campagnolo

@Kev_Shank

@MarkCarey93 & @Mladen_Sormaz

Follow them?

State of the Stats

In order to define the football analytics community, the main reference is Thom Lawrence blog and his State of the Stats survey, published since 2015. From 39 respondents in 2015, the survey attracted more than 200 people in 2016 & widened its Scottish contingent in 2017.

<https://deepxg.com>

Hackathons

Other than the twitter-sphere, an alternative way of accessing this fledgling community is to attend special events such as Hackathons or Data Analytics Fora. Our data collection consists of 24 hours of participant observation at the FIGC Football Data Hackathon in October 2017 and 12 interviews taking place between October and February 2018 when we attended the OptaPro Analytics Forum in London.



Quant vs Real Football Men

Our interest lies in describing the progressive introduction of data science in football coaching and managing. Anecdotal accounts present the 'quants' opposed to traditionalists who do not believe that lessons learnt from stop-start sport can apply to the fluidity of football.

Prozone vs Harry Redknapp

WIRED

Technology

Science

Culture

Gear

Business

Politics

More ▼

When Simon Wilson first arrived at Southampton football club he was a consultant for a technology startup called Prozone. Prozone had developed a proprietary player-tracking software which, fed by eight cameras around the pitch, would output a two-dimensional bird's-eye animation of a football match. The machine could track each player's movement every 0.1 seconds, registering an average of 3,000 touches of the ball per game, and provide an answer to a range of statistical questions. Southampton adopted Prozone and later hired Wilson to work as a performance analyst for the first-team manager.

"Prozone wasn't part of the culture of the game and most managers weren't used to it," Wilson says. "I was naïve but I couldn't understand why they didn't want this kind of information." Once, just before an August 2005 football league Championship game between Luton Town and Southampton, Wilson gave a pre-match briefing to the team and the manager, at the time, Harry Redknapp. "Harry was more intuitive than analytical," says Wilson. "He was nervous about overloading the players with information."

Southampton lost 3-2. On the team bus, Redknapp turned to Wilson and said, "I'll tell you what, next week, why don't we get your computer to play against their computer and see who wins?"

Multiple forms of expertise

We want to understand the process whereby a position is being negotiated for this new form of knowledge and its relation with a much wider array of existing forms of expertise (the coaching, the match analyst, the fitness, the tactical expertise).

Harry Collins on refereeing

Epistemological privilege

Ontological authority

collins, H. (2010) The Philosophy of Umpiring and the Introduction of Decision-Aid Technology, Journal of the Philosophy of Sport, 37, 135-146

“Bad Call”

Collins' programme is to look at the expansion of expertise to include the public. Quite in the opposite direction, we explore how the increasing depiction of football as a 'number game' brings science to bear in a territory where strong is the belief that football is a knowledge acquired by belonging. The eye that understand football is borne out of a body built through long, patient, hard and laborious training.

The changing notion of objectivity

How does 'digital objectivity' (Beaulieu, 2001) evolves vis à vis to the 'trained judgement' (Daston & Gallison, 2007) type of objectivity?

A unique field

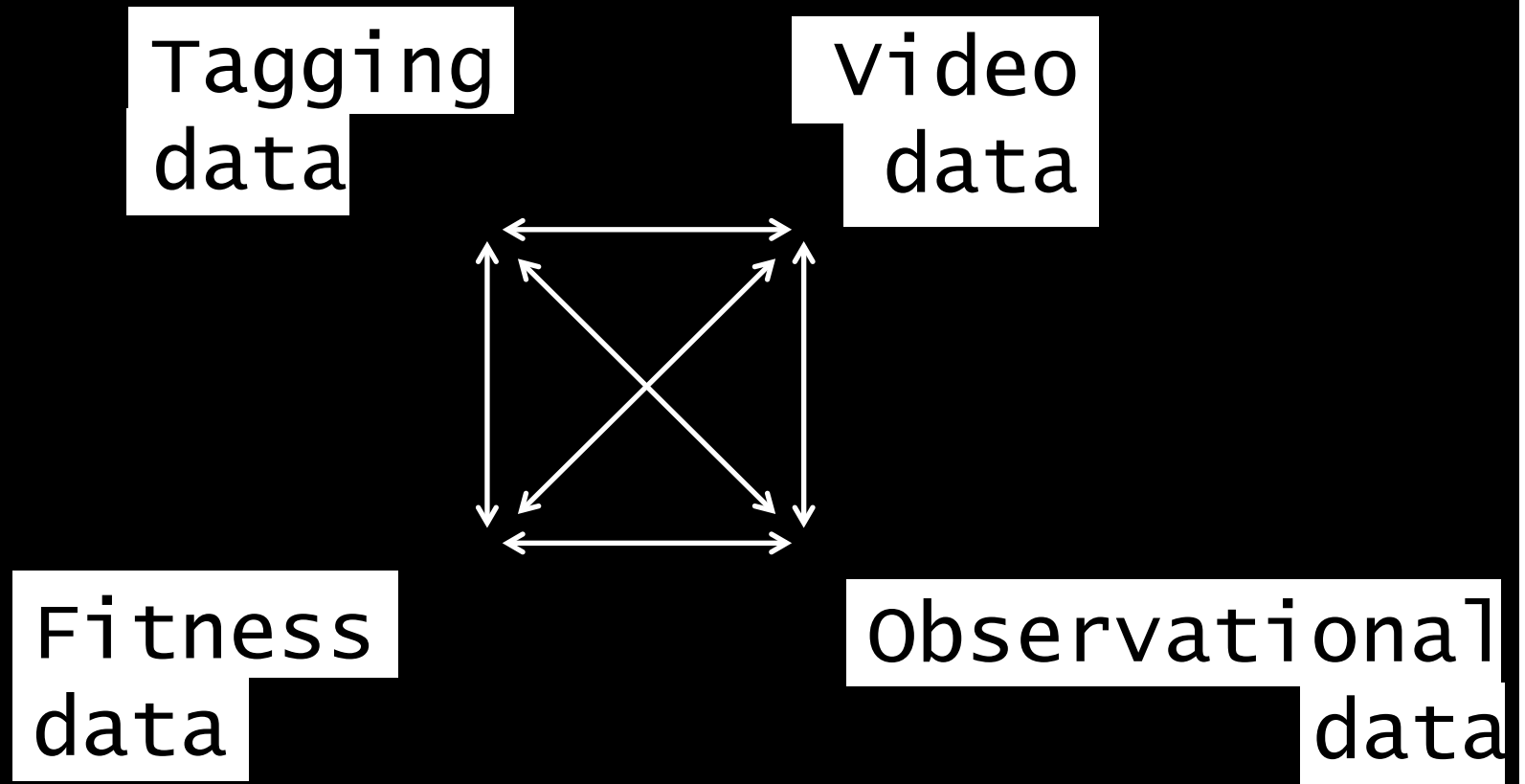
Limited access to data

Star users

speed of technological innovation

resistance to scientific approach

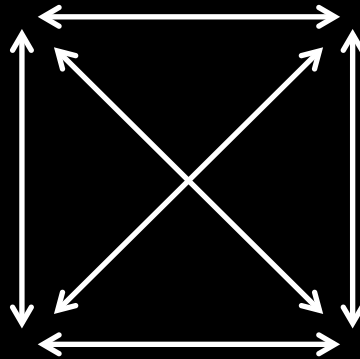
Multiple forms of data



Multiple forms of expertise

Data
Scientist

Match
Analyst



Fitness
Coach

Football
Manager

The Caste

“The world of football is very self-referential. For those who have played in Serie A or B anyone who hasn't played at the same level cannot dare to talk about football. [...] Arrigo Sacchi used to say: “You do not need to have been a horse to be an excellent jockey”” (A.G. Match Analyst)

Primacy of the coach

“L'allenatore é sempre il comandante supremo di tutta la barca/The coach is always the chief commander of the ship” (A.G. Match Analyst)

Equal standing

“The skill of this new profile [the data analyst, ed.] is to understand the today there are many tools available in football. All are useful but none will dominate the others” (D.T. Opta analyst)

Reciprocity

“I always say it is about helping each other. The data will never replace my idea of football or my impressions when watching a match or a player” (D.T. Opta analyst)

Skimming through

The match analyst skims through the numbers and the video to analyse the most important characteristics of our team or the opponent. This is then offered to the coach in the form of a report. There is work a-priori. (D.T. Opta analyst)

Before and after

“In the work of match analysis the ‘live’ is important but even more important are the pre- and the post. There is when you can achieve the maximum degree of accuracy” (D.T. Opta analyst)

Millenials

video data has been somehow legitimized also because all the players these years are twenty years old boys and for the new generations the news come nearly only through images. There is no problem with videos. Data, statistical data, that is difficult it goes to players. We are just starting to pass over to them some tables with performance data.” (A.G. Match Analyst)

Maths vs tactical nous

“Teams from the UK they generally have a greater mathematical approach to the problem...maybe less tactical knowledge. [...] In their team they were all mathematicians. They might have done a great job from the mathematical perspective but from the tactical point of view their interpretations were not so interesting”. (A.G., Match Analyst)

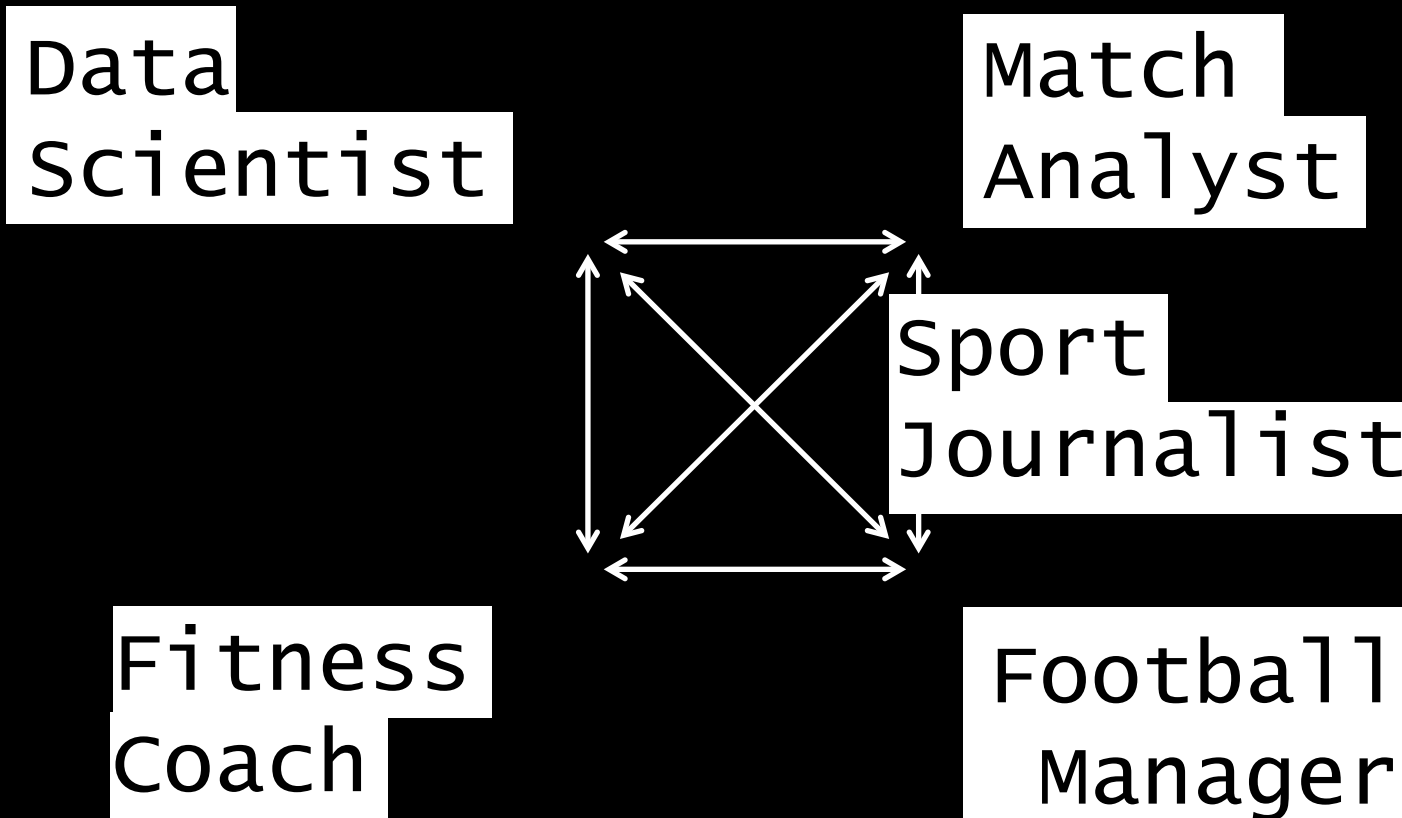
Mixing the data

“Once the technical data [how many shots, how much it ran, how many passes, ed.] is analysed, then it can be mixed and interpreted with other data that traditionally seemed less important [fitness data, ed.]. [...] We can get data out of different sources like medical training or match data and start creating intelligent KPI out of many areas.” (L.S. SAP)

Everything starts with data

“Everything comes from the data. Actually then there is all another subjective part but it's reducing because performance is objective also in sport.” (L.S. SAP)

Multiple forms of expertise



A psychological crutch

“a tool that by its nature is supposed to focus the mind on the factual becomes for many a psychological crutch” (J. Wilson, Guardian)

Summary

Which form of data offers stronger epistemological privilege?

How this translate in acquiring ontological authority?